APPARATUS AND METHOD FOR ABSTRACTING MOTION PICTURE SHAPE DESCRIPTOR INCLUDING STATISTICAL CHARACTERISTICS OF STILL PICTURE SHAPE DESCRIPTOR, AND VIDEO INDEXING SYSTEM AND METHOD USING THE SAME

Technical Field

The present invention relates to an apparatus and method for abstracting motion picture shape descriptors having statistical characteristics of still picture shape descriptors, a video indexing apparatus using the motion picture shape descriptor abstracting apparatus and method, and a computer-readable recording medium for recording a program that implements the motion picture shape descriptor abstracting method.

Background Art

Increasing amount of video and audio data calls for technologies for retrieving and managing the data efficiently. One of these technologies is a multimedia indexing technique for abstracting indexing information representing multimedia data to be used for data retrieval and searching.

Currently, with respect to a still picture, color histograms, shape descriptors and/or texture descriptors are used to abstract indexing information representing multimedia data, and for audio data, spectrum descriptors are used. With respect to a motion picture, motion information descriptors using motion vectors and/or orbit descriptors of objects are used. However, the descriptors of the conventional technologies are not those descriptors used to dynamically index shape information of objects within a video.

In addition, as one dynamic indexing method for indexing the dynamic change of shape data, there is a

method indexing the shape information of an object from the entire still pictures that compose a motion picture or from some representative still pictures by using the conventional still picture shape information indexing

5   method. However, this method has a shortcoming that the data storing and retrieving efficiency is poor, because the amount of indexing information is increased, as the number of still pictures used for abstracting shape data is increased.

10

Disclosure of Invention

It is, therefore, an object of the present invention to provide an apparatus and method for abstracting motion

15 picture shape descriptors by abstracting still picture shape descriptors from the still pictures of an object that compose a motion picture and abstracting motion picture shape descriptors having statistical characteristics from the abstracted still picture shape descriptors to use them

20 as video indexing information, a video indexing system using the motion picture shape descriptor abstracting apparatus and method, and a computer-readable recording medium for recording a program that implements the motion picture shape descriptor abstracting method.

25 In accordance with one aspect of the present invention, there is provided a system for retrieving motion picture, comprising: a motion picture segmentation means for segmenting motion picture temporally; a motion picture shape descriptor abstracting means for abstracting a motion

30 picture shape descriptor from the segmented motion picture data; and a motion picture metadata storing means for storing the motion picture shape descriptor as metadata.

Brief Description of Drawings

35

The above and other objects and features of the

2

present invention will become apparent from the following description of the preferred embodiments given in conjunction with the accompanying drawings, in which:

5   Fig. 1 is a block diagram showing a motion picture shape descriptor apparatus and a motion picture retrieving system in accordance with an embodiment of the present invention;

Fig. 2 is a block diagram illustrating the motion picture shape descriptor abstracter of Fig. 1 in accordance

10   with the embodiment of the present invention;

Fig. 3 is a table showing the metadata stored in a motion picture metadata database for storing motion picture shape descriptors in accordance with the embodiment of the present invention; and

15   Fig. 4 is a flow chart describing a method for abstracting motion picture shape descriptors in accordance with the embodiment of the present invention.

Best Mode for Carrying Out the Invention

20

Generally, shape descriptors of objects for a still picture include outline-based shape descriptors and region-based shape descriptors. The present invention suggests 'a motion picture shape descriptor,' which refers to a

25   descriptor obtained by abstracting shape descriptors, including outline-based shape descriptors or region-based shape descriptors, from the respective still pictures of objects composing a motion picture, and processing the abstracted shape descriptors statistically. The

30   statistically processed motion picture shape descriptors, i.e., statistical characteristic descriptors, have moment characteristics, such as mean and variance.

Following is a process for abstracting motion picture shape descriptors in a statistical shape vector descriptor

35   abstracter.

The shape sequence of an input object is expressed as

3

SS=$\{s_1, s_2, s_3, ..., s_n\}$. Here, s_m denotes an mth shape. A sequence of still picture shape descriptor SD=$\{sd_1, sd_2, sd_3, ..., sd_n\}$ is obtained with respect to each shape from the above shape sequence by using the conventional still picture shape descriptors, e.g., region-based ones or outline-based ones. Here, sd_m is a still picture shape descriptor abstracted from an mth shape s_m. The still picture shape descriptor sd_m is generally expressed as an equation in the form of a vector, i.e., Equation 1 below.

$$sd_m = \{sd_m(1), sd_m(2), sd_m(3), ..., sd_m(1)\}$$

$$\text{Eq. 1}$$

wherein 1 denotes the number of elements that form the vector, and $sd_m(p)$ represents a pth element.

The present invention forms a motion picture shape descriptor by using the sequence SD of the still picture shape descriptor and abstracting four shape descriptors (1) to (4), enumerated below.

(1) Mean Shape Descriptor

Mean shape descriptor $sd^{av} = \{sd^{av}(1), sd^{av}(2), sd^{av}(3), ..., sd^{av}(1)\}$ is abstracted as follows. An mth element $sd^{av}(m)$ is the mean value of the mth element of each of n number of shape descriptors that forms SD = $\{sd1, sd2, sd3, ..., sdn\}$. It can be obtained based on Equation 2.

$$sd^{av}(m) = (\sum_{i=1 \text{ to } n} sd_i(m))/n \qquad \text{Eq. 2}$$

(2) Variance Shape Descriptor

The variance shape descriptor $sd^{var} = \{sd^{var}(1), sd^{var}(2), sd^{var}(3), ..., sd^{var}(1)\}$ is abstracted as follows. That is, an mth element $sd^{var}(m)$ is a variance value of the mth element of each of n number of shape descriptors that form SD = $\{sd1, sd2, sd3, ..., sdn\}$. It can be obtained based on

Equation 3.

$$sd^{var}(m) = (\sum_{i=1 \text{ to } n} (sd_i(m) - sd^{av}(m))^2)/n/(n-1)$$

Eq. 3

(3) Standard Deviation Shape Descriptor

The standard deviation shape descriptor sd^{std} = {$sd^{std}(1)$, $sd^{std}(2)$, $sd^{std}(3)$,..., $sd^{std}(l)$} is abstracted as follows. That is, an mth element $sd^{std}(m)$ is a standard deviation value of the mth element of each of n number of shape descriptors that form SD = {$sd1$, $sd2$, $sd3$,..., sdn}. It can be obtained based on Equation 4.

$$sd^{std}(m) = sqrt(\sum_{i=1 \text{ to } n} (sd_i(m) - sd^{av}(m))2)/(n-1)$$

Eq. 4

(4) Differential Shape Descriptor

The differential shape descriptor shows the change of two consecutive shape descriptors in a shape descriptor sequence. The differential shape descriptor sequence DSD = {dsd_1, dsd_2, dsd_3,..., dsd_{n-1}} can be obtained from the shape descriptor SD = {$sd1$, $sd2$, $sd3$,..., sdn} based on Equation 5.

$$dsd_r = (sd_{r+1} * p_{r+1})(sd_r * p_r)$$

Eq. 5

wherein r is in the range of $0 < r < n$, and p_r denotes a weight of an rth shape descriptor sd_r, which can be obtained from a time rate of a shape represented by a shape descriptor occupying in the entire shape sequence.

The mean shape descriptor, variance shape descriptor and standard deviation shape descriptor, i.e., (1), (2) and (3), are obtained from the differential shape descriptor sequence DSD = {dsd_1, dsd_2, dsd_3,..., dsd_{n-1}} and used for

abstracting motion picture shape descriptors.

The motion picture shape descriptors, suggested in the present invention can use the above shape descriptors alone or a combination thereof. The motion picture shape descriptors abstracted by using a combination of the shape descriptors can be expressed as:

$$CSSD = \{cssd_1, cssd_2, cssd_i, \ldots, cssd_1\}.$$

wherein $cssd_1$ is one of motion picture shape descriptors suggested in the present invention.

If a still picture shape descriptor which is irrespective of the change in size and rotation is applied, a motion picture shape descriptor also irrespective of the change in size and rotation is obtained.

This method of abstracting a statistically processed motion picture shape descriptor can be used with respect to other still picture descriptors, such as still picture texture descriptor, other than the shape descriptors used in the embodiment of the present invention. Therefore, the technology of the present invention has an advantage that it can be generalized.

Other objects and aspects of the invention will become apparent from the following description of the embodiments with reference to the accompanying drawings, which is set forth hereinafter.

Fig. 1 is a block diagram showing a motion picture shape descriptor apparatus and a motion picture retrieving system in accordance with an embodiment of the present invention.

As described in Fig. 1, the motion picture retrieving system includes: a first motion picture shape descriptor abstracting unit 130, a motion picture retrieving device 110, a motion picture database (DB) 120 and a motion

picture shape descriptor metadata DB 150. The motion picture retrieving device 110 includes a second motion picture shape descriptor abstracting unit 130a, a motion picture shape descriptor similarity computing unit 111 and
5 a distance-based classification unit 112.

Hereinafter, the operation of each element will be described.

When segmented motion picture 120 is inputted by a user, the motion picture shape descriptors for the
10 segmented motion picture 120 are abstracted. The abstracted motion picture shape descriptors are inputted to the motion picture shape descriptor similarity computing unit of the motion picture retrieving device 110.

The motion picture stored in the motion picture DB 120
15 for storing motion pictures is inputted to the second motion picture shape descriptor abstracting unit 130a in the motion picture retrieving device 110. Then, the information outputted from the second motion picture shape descriptor abstracting unit 130a is stored in the motion
20 picture shape descriptor metadata DB 150 in the form of metadata. The motion picture shape descriptor similarity computing unit 111 calculates the difference (i.e., similarity) between the motion picture shape descriptors outputted from the first motion picture shape descriptor
25 abstracting unit 130 and the motion picture shape descriptors in the motion picture shape descriptor metadata DB 150. To calculate the similarity (i.e., distance), a method using Euclidian distance which measures the distance between two vectors or a method using the sum of absolute
30 difference is used. The distance-based classification unit 112 sorts out the calculated distance information in the order of distances from close to far, abstracts corresponding metadata information from the motion picture shape descriptor metadata DB 150, and outputs the
35 abstracted similar motion picture information 140 to the user.

Fig. 2 is a block diagram illustrating the motion picture shape descriptor abstracter of Fig. 1 in accordance with the embodiment of the present invention. As illustrated in the drawing, the motion picture shape
5   descriptor abstracting unit 230 of the present invention includes: a motion picture segmentation unit 210, a motion picture shape descriptor abstracting unit 230 and a motion picture metadata DB 250. The motion picture shape descriptor abstracting unit 230 includes a shape abstracter
10 231, a shape vector descriptor abstracter 233 and a statistical shape vector descriptor abstracter 235.

Hereinafter, the operation of each element will be described. First, a motion picture 200 is inputted to the motion picture segmentation unit 210 and segmented
15 temporally. The temporally segmented motion picture 200 is inputted to the shape abstracter 231, which then outputs shape information motion picture 232, corresponding to one object. The shape information of each still picture of the shape information motion picture 232 is inputted to the
20 shape vector descriptor abstracter 233, which outputs a shape vector descriptor sequence 234.

The shape vector descriptor sequence 234 outputted from the shape vector descriptor abstracter 233 is inputted to the statistical shape vector descriptor abstracter 235.
25 The statistical shape vector descriptor abstracter 235 outputs a motion picture shape descriptor 240 eventually, by using each or a combination of the Equations 1 through 5, each of which corresponds to the above enumerated (1) mean shape descriptor, (2) variance shape descriptor, (3)
30 standard deviation shape descriptor and (4) differential shape descriptor. The motion picture shape descriptor 240 is stored in the motion picture metadata DB 250 for storing motion picture metadata.

Fig. 3 is a table showing the metadata stored in a
35 motion picture metadata database for storing motion picture shape descriptors in accordance with the embodiment of the

present invention. The metadata are classified based on motion picture shape vector descriptor, motion picture title, location of file, and the location of starting time in the original motion picture.

5   Fig. 4 is a flow chart describing a method for abstracting a motion picture shape descriptor in accordance with the embodiment of the present invention. As shown in the drawing, to abstract a motion picture shape descriptor, at step S403, an input motion picture 400 is segmented
10   temporally, and at step S405, a shape information of motion picture corresponding to one object is abstracted from the temporally segmented motion picture.

Subsequently, at step S407, a shape vector descriptor sequence is abstracted from the abstracted shape
15   information of motion picture. At step S409, a motion picture shape descriptor, which is a statistical shape descriptor, is abstracted from the shape vector descriptor sequence. Then, at step S411, the abstracted motion picture shape descriptor is stored in the motion picture
20   metadata DB for storing motion picture metadata.

As described above, the technology of the present invention can store the changing shape information of a motion picture object effectively by using a motion picture shape descriptor, and using the stored motion picture
25   information for retrieving motion picture and, further, for video indexing.

While the present invention has been described with respect to certain preferred embodiments, it will be apparent to those skilled in the art that various changes
30   and modifications may be made without departing from the scope of the invention as defined in the following claims.